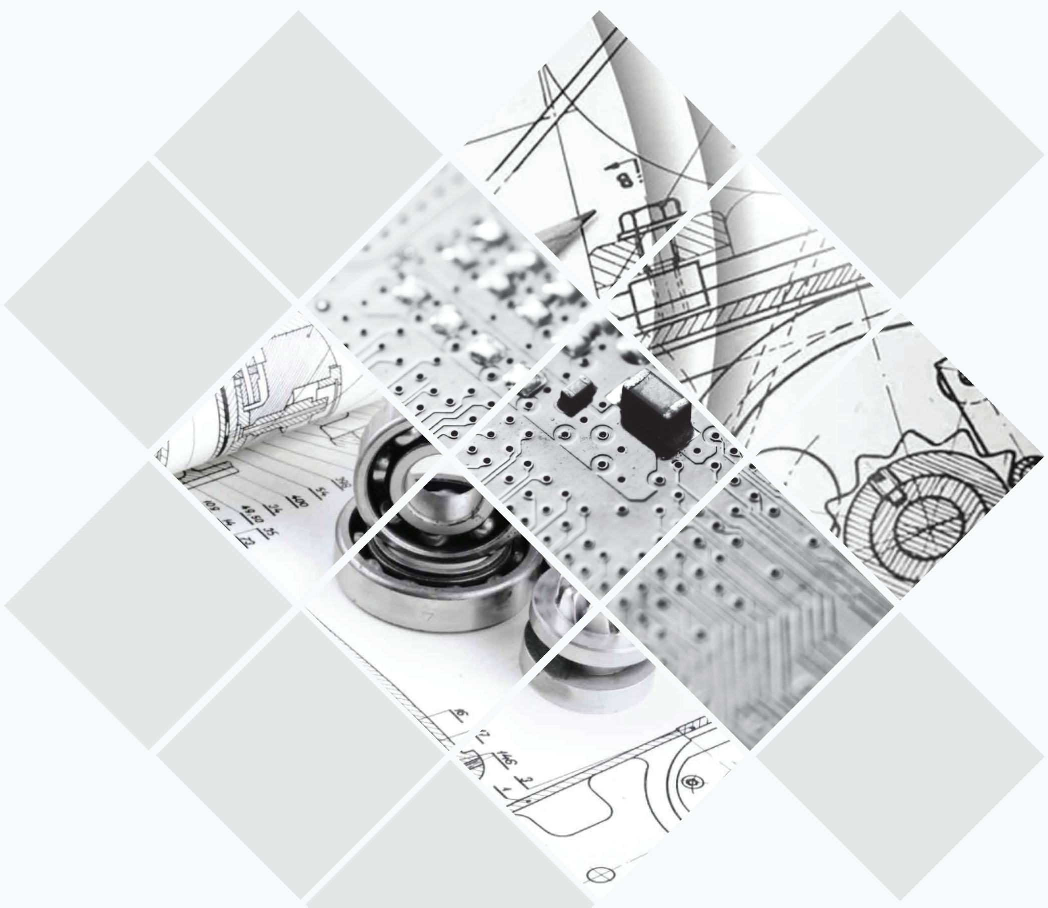
BORN TO ENGINEER

test optimization using adaptive random testing techniques

In today's competitive software development scenario, the customer demands a testing coverage which not only ensures the stated requirements but also the implied ones.



S Prem Sagar
Indumol KS
Arunkumar V

QuEST Global

contents

Test Optimization Using Adaptive Random Testing Techniques

ABSTRACT

In today's competitive software development scenario, the customer demands a testing coverage which not only ensures the stated requirements but also the implied ones. This situation calls for an exhaustive testing which may not be always possible due to various reasons. Testing, due to its last position in SDLC, often gets crunched due to the cumulative schedule slippages. Hence Tester is faced with a challenge to make testing as efficient as possible within a short time span due to cost constraints. With selective testing an only option, test leads usually go for the age-old approach of Random Testing. Random testing does not ensure coverage in a scientific manner. Hence the quality and coverage of the testing cannot be guaranteed.

This paper explains the concept of Adaptive Random Testing in which each test case is selected such that it covers a scenario which would be dimensionally farthest from the previously conducted test.

Index Terms – Random Testing, Anti-Random Testing, Adaptive Random Testing, ART Series, ART for Binary, Cartesian Distance, Hamming Distance.

I. INTRODUCTION

EXHAUSTIVE testing may not be always possible in software development life cycles due to various reasons like budget limitation, schedule crunches etc. Since exhaustive testing will turn out to be an expensive agenda, managements are on a lookout for alternatives.

Though Random Testing is often used as a cost reduction strategy wherein testing a product is done using test cases and test data that has been chosen at random, it may ensure a satisfying test coverage due to its unscientific nature Although the judgment of the test lead or test engineers is used for choosing the test cases, indicative attributes like the test cases previously executed are not considered.

While methods like Taguchi, pair wise testing etc lend a technical approach to defining efficient test cases, these methods do not address the optimal sequence for executing the

test cases. The usage of numerical strategies to identify the sequence of tests giving maximum coverage would lend more credibility to the final testing by increasing the probability of detecting failures.

The objective of this paper is to explain the Adaptive Random Testing (ART) technique. For a set of suitably designed tests, this strategy helps to enhance the testing process by

• Defining an optimal sequence of test execution.

• Defining the point beyond which executing of further test cases may not be beneficial (for extremely short and critical testing phases).

• Selecting critical scenarios for execution in case of multiple environments.

Test Optimization Using Adaptive Random Testing Techniques

II. ADAPTIVE RANDOM TESTING

Adaptive random testing uses mechanisms to ensure much even and widespread distributions of test cases over an input domain. It is an effective improvement of Random Testing (RT) in the sense that fewer test cases are needed to detect the first failure. It is based on the observation that failure-causing inputs are normally clustered in one or more contiguous regions in the input domain. Hence, the test execution should refer to the locations of successful test cases (those that do not reveal failures) to ensure that all test cases are far apart and evenly spread in the input domain.[3] i.e., for a function f(p,q), random testing would choose p and q in an ad-hoc fashion while adaptive random testing attempts to choose p and q such that maximum coverage of input domain is obtained in minimal time.

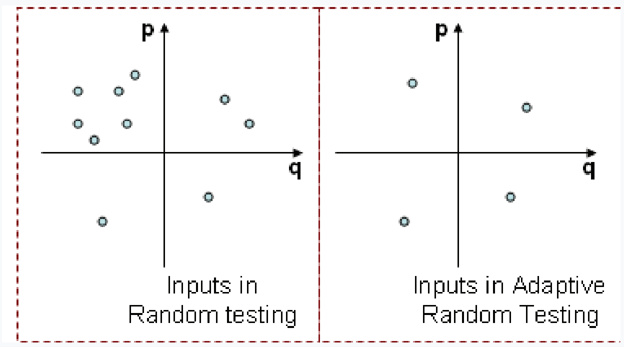


Figure 1: Input Selection

In ART, new tests are selected religiously by ensuring the distance of the new test to be maximum from all the previously run tests. This ensures that every new test conducted will execute some new lines of code which are not executed by the previous tests. The selected test will then have maximum probability of exposing a non-captured bug.

Consider the case of testing a 3-dimensional cube.

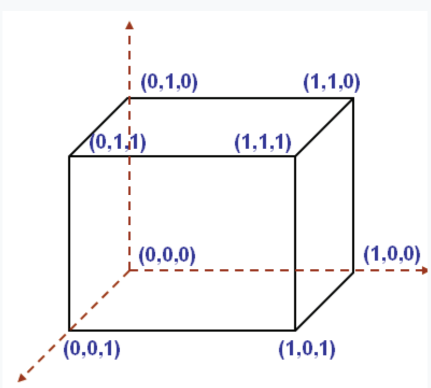


Figure 2: 3-Dimensional Cube for Testing

If you start with testing the corner with Cartesian coordinates (0,0,0), to ensure maximum coverage, the next test should be for the corner with Cartesian coordinates (1,1,1). This is the basic concept behind ART.

The distance can be computed either by Hamming Distance or Cartesian Distance methods.

A. Hamming Distance

Hamming distance is the number of bits in which two binary vectors differ. It is not defined for vectors containing continuous value. Total Hamming distance for any vector is the sum of its Hamming distance with respect to all the previous vectors

For e.g. Hamming distance between

- **1011101 and 1001001 is 2.**
- **2173896 and 2233796 is 3.**
- **"toned" and "roses" is 3.**

B. Cartesian Distance

Cartesian distance between two vectors

A = {a0, a1…, an} and B = {b0, b1…, bn} is given by

CD (A, B) = [(a0-b0)2 + (a1-b1)2 + … + (an-bn)2

Total Cartesian distance for any vector is the sum of its Cartesian distance with respect to all previous vectors. It can also be seen that if all the variables in two vectors are binary, then

CD (A, B) = HD (A, B)

For e.g., consider the example of finding distance between two vectors A = {0,0,1,0,1} and B = {1,1,0,0,1}

HD (A, B) = | 0 – 1| + | 0 – 1| + | 1- 0 | + | 0 – 0 | + | 1 – 1 | = 3

CD (A, B) = [(0-1)2 + (0-1)2 + (1-0)2 + (0-0)2 + (1-1)2] = 3

C. Maximal Distance Adaptive Random Testing Sequence

Maximal Distance ART Sequence is a test sequence such that a test 'ti' is chosen such that the distance of 'ti' from all the previously conducted tests t0, t1,…, ti-1 is maximum when compared with all the remaining test scenarios ti+1, …, tn.

i.e. Total Distance, TD (ti) = j=0 i-1 D (ti, tj); D (ti, tj) indicates the distance between the vectors ti and tj.

Test Optimization Using Adaptive Random Testing Techniques

D. Generation of ART Sequence

To understand the concept of determining distances in the input domain, we go through the generation of some basic adaptive random sequences. We are considering only the Cartesian Distance for finding the distance between two vectors.

E. ART for 4 factors taking binary values

Consider a function which is having four factors w, x, y and z is to be tested. All the factors can take only binary values i.e. either 0 or 1. Totally there will be 24 = 16 test scenarios as shown in Table 1.

TABLE 1

ALL POSSIBLE COMBINATIONS OF TEST SCENARIOS FOR A 4 FACTOR BINARY CASE

	w	x	y	z
vector 1	0	0	0	0
vector 2	0	0	0	1
vector 3	0	0	1	0
vector 4	0	1	0	0
vector 5	1	0	0	0
vector 6	1	1	0	0
vector 7	1	0	1	0
vector 8	1	0	0	1
vector 9	0	1	1	0
vector 10	0	1	0	1
vector 11	0	0	1	1
vector 12	0	1	1	1
vector 13	1	0	1	1
vector 14	1	1	0	1
vector 15	1	1	1	0
vector 16	1	1	1	1

We need to identify the sequence of test scenarios which will give us the maximum coverage when executed. Without losing the generality we can choose the first vector for testing as A1 = {0, 0, 0, 0}. Now we need to find the test scenario, which is at a maximum distance from the selected scenario. Taking Cartesian distance between the vectors, from Table 2, it can be seen that Vector 16 is the vector which is at a maximum distance from A1.

TABLE 2

DISTANCE FROM VECTOR A1 TO OTHER VECTORS

	w	x	y	z	$D(t_i,t_0)$
A_1	0	0	0	0	
vector 2	0	0	0	1	1.00
vector 3	0	0	1	0	1.00
vector 4	0	0	1	1	1.41
vector 5	0	1	0	0	1.00
vector 6	0	1	0	1	1.41
vector 7	0	1	1	0	1.41
vector 8	0	1	1	1	1.73
vector 9	1	0	0	0	1.00
vector 10	1	0	0	1	1.41
vector 11	1	0	1	0	1.41
vector 12	1	0	1	1	1.73
vector 13	1	1	0	0	1.41
vector 14	1	1	0	1	1.73
vector 15	1	1	1	0	1.73
vector 16	1	1	1	1	2.00

Vector 16 can be selected as Select 2. i.e. A2 = {1, 1, 1, 1}. Now the select 3 has to be taken such that it is at a maximum distance from the already selected two vectors. From Table 3, indicates that there are 6 vectors which are equidistant from A1 and A2. Without losing generality we can choose one vector from the available 6 vectors. Let us choose A3 = {0, 0, 1, 1}.

Test Optimization Using Adaptive Random Testing Techniques

TABLE 3
DISTANCE FROM A1 AND A2 TO OTHER VECTORS

	w	x	y	z	$D(t_i,t_0)$
A_1	0	0	0	0	
A_2	1	1	1	1	
vector 2	0	0	0	1	2.73
vector 3	0	0	1	0	2.73
vector 4	0	0	1	1	2.83
vector 5	0	1	0	0	2.73
vector 6	0	1	0	1	2.83
vector 7	0	1	1	0	2.83
vector 8	0	1	1	1	2.73
vector 9	1	0	0	0	2.73
vector 10	1	0	0	1	2.83
vector 11	1	0	1	0	2.83
vector 12	1	0	1	1	2.73
vector 13	1	1	0	0	2.83
vector 14	1	1	0	1	2.73
vector 15	1	1	1	0	2.73

From Table 4 we identify that the next vector at maximum distance from A1, A2 and A3 is vector 13 and so A4 will be {1, 1, 0, 0}.

TABLE 4
DISTANCE FROM VECTOR A1, A2 AND A3 TO OTHER VECTORS

	w	x	y	z	$D(t_i,t_0)$
A_1	0	0	0	0	
A_2	1	1	1	1	
A_3	0	0	1	1	
vector 2	0	0	0	1	3.73
vector 3	0	0	1	0	3.73
vector 5	0	1	0	0	4.46
vector 6	0	1	0	1	4.24
vector 7	0	1	1	0	4.24
vector 8	0	1	1	1	3.73
vector 9	1	0	0	0	4.46
vector 10	1	0	0	1	4.24
vector 11	1	0	1	0	4.24
vector 12	1	0	1	1	3.73
vector 13	1	1	0	0	4.83
vector 14	1	1	0	1	4.46
vector 15	1	1	1	0	4.46

In a similar fashion, the rest of the ordering can be computed and the resulting adaptive random test sequence will be as shown in Table 5.

TABLE 5
ADAPTIVE RANDOM TEST SEQUENCE

	w	x	y	z
A_1	0	0	0	0
A_2	1	1	1	1
A_3	0	0	1	1
A_4	1	1	0	0
A_5	0	1	0	1
A_6	1	0	1	0
A_7	1	0	0	1
A_8	0	1	1	0
A_9	0	0	0	1
A_{10}	1	1	1	0
A_{11}	0	0	1	0
A_{12}	1	1	0	1
A_{13}	0	1	0	0
A_{14}	1	0	1	1
A_{15}	0	1	1	1
A_{16}	1	0	0	0

III. ART IN PROJECTS

Generation of the adaptive random testing sequence includes the following steps.

A. Identification of the test situation

This includes analyzing the system under test. The factors based on which tests should be designed are to be identified. This could include the test environments as well as test data. Two possible scenarios for where ART can be used includes

1) Configuration Testing:

Internationalization of software products being a key feature today, configuration testing is essential for software products which are expected to work in on a wide number of platforms and operating systems. The various conditions including operating system, language etc can be considered as different factors. ART is done for selecting critical test environments in a product expected to work in 16 languages and 7 OS.

2) User Interface Testing:

Widespread usage of software applications like online forms opens the user interfaces to a wide and often untrained mass of people. Hence the testing of these user interfaces becomes very critical. The various fields with their possible entries can be considered as different factors

Test Optimization Using Adaptive Random Testing Techniques

2) User Interface Testing:

Widespread usage of software applications like online forms opens the user interfaces to a wide and often untrained mass of people. Hence the testing of these user interfaces becomes very critical. The various fields with their possible entries can be considered as different factors

B. Designing for ART

It is not practical to assume that all systems under test will have only factors with binary inputs. The systems under test need to be mapped to numerical terms to effectively use ART.

1) Case 1: non-binary two factor case:

To test the form shown in Fig 3, with 3 factors each taking two values, the factors are converted to a binary set using the guideline in Table 6.

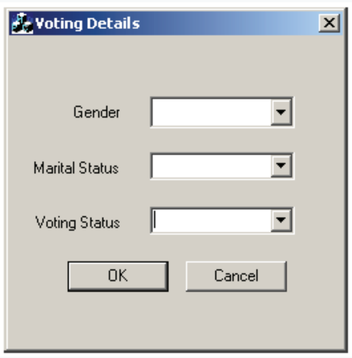


Fig. 3. FORM FOR VOTING DETAILS

TABLE 6

MAPPING OF VARIABLE FACTORS TO BINARY TERMS

Factor/Value	0	1
Gender	Male	Female
Marital Status	Single	Married
Voting Status	Yes	No

Now the problem reduces to a simple ART sequence generation problem. Hence vectors are defined as A1 = {0, 0, 0}; A1 = {Male, Single, Yes}, A2 = {1, 1, 1} A2 = {Female, Married, No}.etc

2) Case 2: Non-Binary More than Two Factor Case:

Consider a software application involving a database of employees of different levels. Let the factors be ‘Gender’ of the employee (Male and Female), ‘Designation’ in which he/she is placed in the organization (SE, SSE and LE) and the ‘Grade’ obtained for each designation (1, 2, 3 and 4). For ART scenario creation, the first step is to replace the non-numerical values into numerical form. As a couple of factors have more than two possible levels of inputs, the resultant mapping will not be of binary form. Table 7 shows the mapping of factors and inputs.

TABLE 6

MAPPING OF VARIABLE FACTORS TO BINARY TERMS

Factor/Value	0	1	2	3
Gender	Male	Female		
Designation	SE	SSE	LE	
Grade	1	2	3	4

For ART, a variation in each of the factors should have the same impact while generating the sequence. But in the above case we can see the factor ‘gender’ can vary only from 0 to 1 while the factor ‘grade’ takes a value from 0 to 3. As they are not assigned values uniformly, with the current assignment, the factor ‘grade’ is having a greater influence in the distance computations, which is illogical.

To overcome the lacuna, we have to normalize the levels of the factors. By normalizing the factors, Table 7 can be rewritten as shown in Table 8. These normalized values can be used for the ART sequence creation.

TABLE 8

NORMALIZED MAPPING OF VARIABLE FACTORS TO BINARY TERMS

Factor/Value	0	0.33	0.5	0.67	1
Gender	Male				Female
Designation	SE		SSE		LE
Grade	1	2		3	4

C. Generating ART

The ART sequences for Case 1 and Case 2 in section 3.2 are as follows:

Test Optimization Using Adaptive Random Testing Techniques

TABLE 9

ART SEQUENCE FOR VOTING APPLICATION

Sl No:	Gender	Marital Status	Voting Status
A_1	Male	Single	Yes
A_2	Female	Married	No
A_3	Male	Single	No
A_4	Female	Married	Yes
A_5	Male	Married	No
A_6	Female	Single	Yes
A_7	Male	Married	Yes
A_8	Female	Single	No

TABLE 10

ART SEQUENCE FOR EMPLOYEE DATABASE

Sl No:	Gender	Designation	Grade
A_1	Male	SE	1
A_2	Female	LE	4
A_3	Male	SE	3
A_4	Female	LE	1
A_5	Male	LE	2
A_6	Female	SSE	4
A_7	Female	SE	1
A_8	Male	SSE	4
A_9	Male	LE	3
A_{10}	Male	SE	4
A_{11}	Male	SE	2
A_{12}	Female	LE	2
A_{13}	Female	SE	3
A_{14}	Female	LE	3
A_{15}	Male	SSE	1
A_{16}	Female	SE	2
A_{17}	Male	LE	1
A_{18}	Female	SE	4
A_{19}	Male	LE	4
A_{20}	Female	SSE	1
A_{21}	Male	SSE	2
A_{22}	Female	SSE	3
A_{23}	Male	SSE	3
A_{24}	Female	SSE	2

IV. ADVANTAGES OF ART SEQUENCES

ART ensures that the test scenarios are executed such that the maximum coverage is obtained with minimum scenarios and the scenarios with maximum coverage are executed first. This helps in early defect finding thereby reducing the effective cycle time.

ART sequence identifies the preferred order of test execution for the exhaustive set of test cases thus providing flexibility for the test lead to identify the test exit point.

V. LIMITATIONS OF ADAPTIVE RANDOM TESTING SEQUENCES

ART is effective when the number of variable and factors increases. But it can be noticed that as the number of variables and factors goes up, it is almost impossible to derive the ARTS manually. Test application has to be made which will generate the ART sequence once the factors and input levels are given.

VI. FUTURE WORK

The examples shown above sequenced all the possible scenarios. In a situation with minimal time for testing, two strategies can be adopted to cut down the number of test scenarios logically.

1) Defining ART Index:

ART Index identifies the number of test scenarios in the ART sequence to be executed in order to attain a defined level of stability. ART Index is defined as a threshold value of average distance between the test scenarios At least all test scenarios with average distance till this defined index should be executed. The average distance can be defined as the average of distances of newly selected vector from all the previous vectors. In the above problem, if the ART Index is defined as 1.4, it indicates that all test scenarios which are at a distance of 1.4 of beyond from the previously executed test scenarios will be executed. The level of stability is inversely proportional to the ART Index.

2) Adopting Taguchi:

Taguchi strategies can be adopted for selecting the critical scenarios to be tested [4]. Determining ART sequence from this will help to ensure that the critical scenarios are run such that maximum coverage is obtained even if the complete sequence cannot be run.

VII. CONCLUSION

Adaptive Random testing is a new test generation approach that requires further theoretical and experimental investigations. Here we have presented its effectiveness to optimize the testing process by keeping tests as different as possible from each other. The limitation of adaptive random testing confining to binary values has been overcome by mapping process. Effective use of ARTS ensures that the faults are identified earlier and chaotic testing phase can be brought under control.

Test Optimization Using Adaptive Random Testing Techniques

IX. REFERENCES

[1] 1. Yashwant K Malaiya, "Adaptive Random Testing : Getting the most out of Black Box Testing", Technical Report 96-129, Computer Science Department, Colorado State University, Fort Collins, CO, 1996

[2] 2. G Meyers, "Art of Software Testing", John Wiley & Sons, 1979

[3] 3. T.Y. Chen, Zongyuan Yang, T.H. Tse and De Hao Huang, An Innovative Approach to Tackling the Boundary Effect in Adaptive Random Testing: Proceedings of the 40th Annual Hawaii International Conference on System Sciences (HICSS-40)

[4] 4. Unni Sankar, Deepa Thampy, Applying Taguchi Methods in Software Product Engineering, Proceedings of QAI's SEPG Conference

X. AUTHOR PROFILE

Indumol K S is a Project Leader at QuEST, Trivandrum. She is a Computer Engineering diploma holder and has an industrial experience of 7 years. She is an ISTQB Certified Test Engineer. She is specialized in Medical & HealthCare domains. She plays an active role in the Independent Testing Team (ITT) of QuEST.

S Prem Sagar is currently working as a Program Manager at QuEST, Technopark Trivandrum. He is a Bachelor of Engineering and has above 10 yrs of IT experience performing Testing & Quality Management, Project Management & Web development. He had previously worked as a Test Lead with Pentasoft Technologies, Tidel Park, Chennai. He is a member of Project Management Institute and a certified ISTQB test engineer.

Arun Kumar V is working as Assistant Manager at QuEST Trivandrum. He did his B-Tech from Government Engineering College, Idukki and has more than 3.5 years of experience in the industry. He played a key role in implementing the models like ISO 27001, ISO9001:2008, ISO 13485. He played a key role in the implementation of high maturity artifacts of CMMI V1.2. He is working as a Process Consultant at NeST. He is the convener of ISF and is responsible for the ISMS related activities in the organization. He is CQSPE (by QAI) certified and an ISTQB Certified Tester.

BORN TO ENGINEER

www.quest-global.com